

Il servizio Gateway HTTP-Z39.50

Daniela Dragone, Paola Rossi

CILEA, Segrate

Abstract

In questo articolo viene presentato il servizio Gateway HTTP-Z39.50 del CILEA. Questo servizio permette ad un utente Internet di interrogare via web gli OPAC (Online Public Access Catalog) dotati di server Z39.50. Tali OPAC usano il protocollo Z39.50 e sono solitamente consultabili utilizzando client specifici. Tramite il Gateway, liberamente disponibile per i gestori degli OPAC Z39.50, è possibile connettere tali OPAC alla metaricerca fornita dal servizio MAI - MetaOPAC Azalai Italiano.

The CILEA Gateway HTTP-Z39.50 is a service that allows an Internet user to locate information in a Z39.50 OPAC (Online Public Access Catalogs) using the web. This kind of OPAC is an application of the standard bibliographic protocol Z39.50 and can generally be searched using specific clients. By the Gateway service, freely available for Z39.50 OPAC administrators, the Z39.50 OPACs can be connected to the metasearch provided by MAI - MetaOPAC Azalai Italiano.

Keywords: Beni culturali, Biblioteche, Biblioteca virtuale, Catalogo, Gateway, Http, MAI, MetaOPAC, OPAC, Z39.50.

Dal mese di aprile 2005 è disponibile per l'utenza Internet, con particolare attenzione alla comunità bibliotecaria, il servizio Gateway HTTP-Z39.50 all'indirizzo <http://gwz.cilea.it>.

Obiettivo di questo servizio è quello di offrire un'interfaccia web di ricerca agli OPAC (*Online Public Access Catalog*, cataloghi bibliografici ad accesso aperto) che utilizzano il protocollo Z39.50, un protocollo standard diffuso nella comunità bibliotecaria.

I servizi OPAC Z39.50 collegati al Gateway possono richiedere la connessione al MAI - MetaOPAC Azalai Italiano, la cui metaricerca interroga attualmente più di duecentoquaranta sistemi OPAC Italiani.

In questo articolo viene brevemente introdotta la problematica relativa agli OPAC Z39.50 e presentato il servizio.

Lo standard Z39.50 e gli OPAC

Storicamente il protocollo Z39.50 è stato realizzato ed usato prima della diffusione del web, principalmente presso le Università americane, molto avanzate nella realizzazione di strumenti per le biblioteche.

Z39.50 è oggi il nome più conosciuto per lo standard ISO 23950: "Information Retrieval (Z39.50): Application Service Definition and

Protocol Specification": tale protocollo viene utilizzato principalmente in ambito bibliotecario e lo standard è gestito oggi dalla Library of Congress [1].

Il protocollo Z39.50 può essere visto a grandi linee come costituito da due componenti.

La prima componente è costituita dallo standard bibliografico di recupero di dati (*retrieval*), mentre la seconda è costituita dalla codifica di comunicazione e trasmissione fra server e client.

Lo standard bibliografico identifica, attraverso una tabella di sei tipi di attributi, l'information retrieval applicabile. Sono indicati gli accessi bibliografici, chiamati *use attributes*, cioè le chiavi d'accesso o indici bibliografici (come l'autore o il titolo, ecc.) che sono disponibili all'utente per interrogare i vari campi bibliografici del catalogo (p.e. la responsabilità d'autore, ecc.). I campi bibliografici sono codificati secondo vari standard di descrizione bibliografica, come ISBD (*International Standard Book Description*), ecc. Gli altri *attributes* indicano poi come sono costruite tecnicamente le ricerche su questi dati testuali (*retrieval* esatto sull'intero campo testuale o solo su una sua componente, con quali

approssimazioni di troncamento, ecc.), indicando, per ognuna delle modalità di match tra dato richiesto e dato registrato, la serie delle possibilità contemplate. Ogni installazione di catalogo poi si caratterizza definendo le proprie opzioni scelte (profilo).

La connessione e trasmissione da e verso un server Z39.50 risultano congeniali al protocollo bibliografico: attraverso l'apertura di una sessione Z39.50, p.e. i risultati delle varie ricerche fatte dall'utente sono recuperabili anche successivamente, incrociabili con altre condizioni anche nei passi successivi della sessione di ricerca.

La sessione Z39.50 rimanda ad una concezione differente da quella, tipica del web, "stateless", che si è andata diffondendo nell'utilizzo della rete Internet.

Il servizio di Gateway proposto facilita l'utente di Internet nella consultazione dei server OPAC Z39.50, lasciando alla risorsa OPAC Z39.50 le sue caratteristiche realizzative.

Normalmente gli OPAC che utilizzano il protocollo Z39.50 sono interrogabili via Internet utilizzando client specifici per il protocollo. Il Gateway HTTP-Z39.50 è stato realizzato allo scopo di consentire ad un utente Internet di interrogare gli OPAC utilizzando i più comuni browser web (Mozilla, Internet Explorer, ecc.) senza che occorra usare client specifici.

I browser utilizzano principalmente il protocollo HTTP (HyperText Transport Protocol) [2] che è quello su cui si basa il World Wide Web e che viene utilizzato per trasmettere e collegare le informazioni in Internet. Sono organizzate in pagine scritte in HTML (HyperText Markup Language) [3], un linguaggio che permette di controllare il formato del testo, di collegare (link) altre pagine, di inserire elementi grafici, audio e video.

Le applicazioni OPAC di tipo web mettono a disposizione proprie pagine web in risposta a richieste specifiche degli utenti attraverso dei moduli di ricerca. Le modalità con cui il sistema

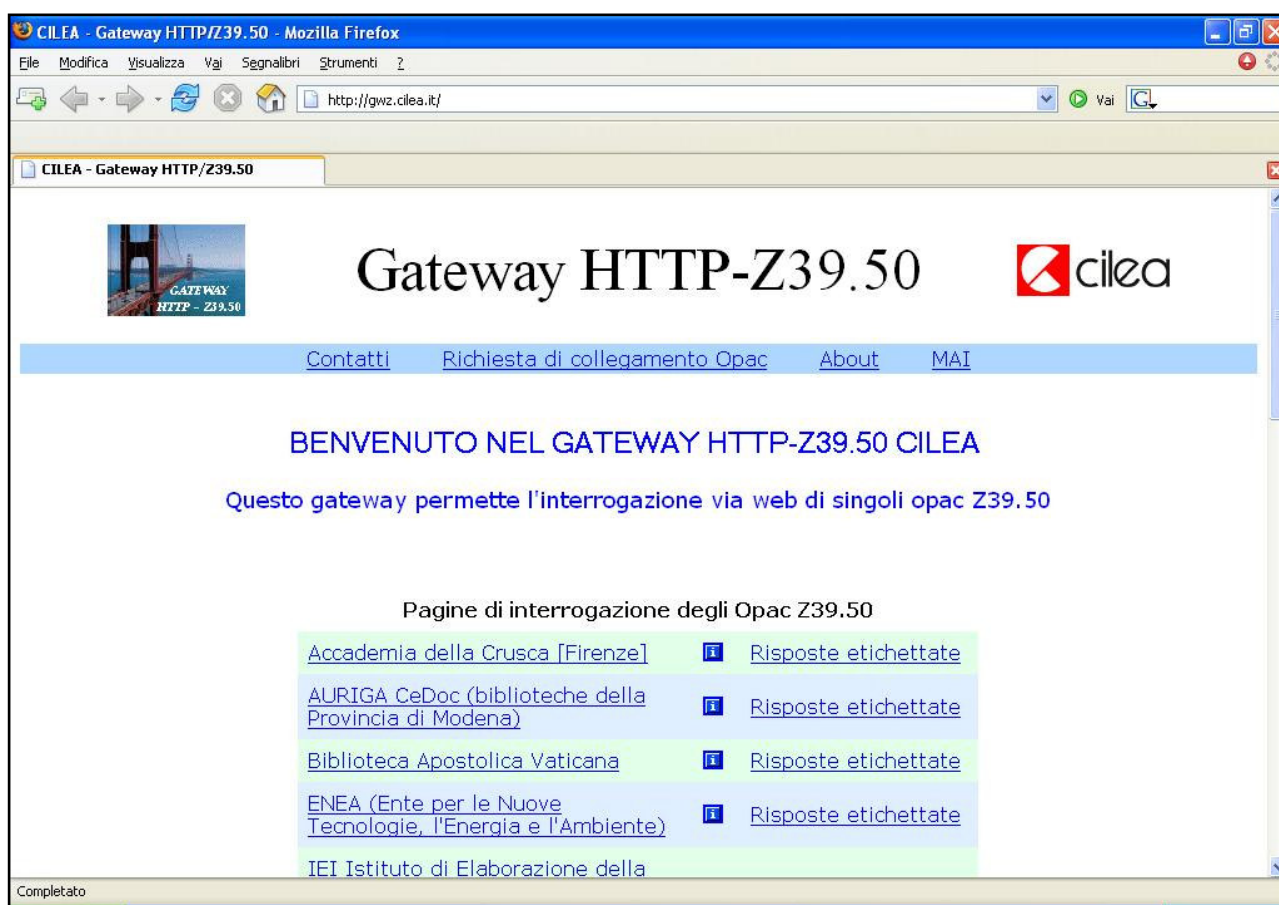


Fig. 1 - Homepage del Gateway

fa incontrare dati e richieste sono dettate dal retrieval.

Il servizio Gateway permette l'interrogazione di OPAC, che utilizzano il protocollo Z39.50, attraverso la compilazione di moduli di ricerca. Tali moduli sono contenuti in semplici pagine web dall'aspetto simile a quelle proposte all'utenza Internet dagli OPAC. D'altro canto, il sistema Z39.50 risponde nella propria modalità testuale e non ipertestuale, cioè senza link informativi bibliografici ulteriori, ed il servizio complessivamente fornito dal Gateway risulterà analogo, senza ulteriore navigazione puntuale e locale sulla banca dati del catalogo, a differenza dell'applicazione OPAC web.

Il servizio è liberamente usufruibile, sia come ricerca bibliografica che come possibilità di connessione per un servizio OPAC Z39.50. Per questo scopo si invitano i gestori a segnalare al gestore del Gateway i dati specifici dei propri servizi Z39.50 attraverso la pagina "Richiesta di collegamento Opac", <http://gwz.cilea.it/collegamento.html>. Dopo aver verificato la completezza dei dati segnalati e valutato l'idoneità dell'OPAC al servizio, il gestore connette l'OPAC e si occupa della corretta configurazione del Gateway. Da questo momento l'OPAC è raggiungibile dalla homepage del servizio.

Fornendo un'interfaccia (pagine web) di facile utilizzo, il Gateway utilizza l'OPAC Z39.50 senza alterare le sue caratteristiche e ne permette una più ampia raggiungibilità via web.

Architettura

Il Gateway è stato realizzato utilizzando il software ZAP [4], un modulo di Apache, un server HTTP di tipo open source [6], e alcuni programmi sviluppati appositamente nel lin-

Il Gateway HTTP-Z39.50:

- riceve richieste in HTTP dal client,
- traduce la richiesta in Z39.50,
- manda la richiesta all'OPAC Z39.50 selezionato,
- riceve la risposta in Z39.50,
- traduce la risposta in HTTP
- manda la risposta al client in HTTP.

L'homepage del Gateway (<http://gwz.cilea.it>) presenta l'elenco degli OPAC interrogabili. Di fianco ad ogni nome di OPAC è presente l'immagine **1** che consente di raggiungere la scheda informativa dell'OPAC derivante dal servizio MAI di Repertorio [5].

Su questa pagina è possibile interrogare l'OPAC utilizzando due link: nome dell'OPAC e risposte etichettate.

La differenza tra tali pagine risiede nella modalità di risposta:

- interrogando l'OPAC tramite il link con il nome dell'OPAC, il sistema visualizzerà i record di risposta nel loro formato nativo di sintassi (es: Unimarc, Sutrs,...), senza alcuna formattazione;
- interrogando l'OPAC tramite il link risposte etichettate, i record di risposta vengono con i campi etichettati (tradotti) in modo da renderli più comprensibili ai meno esperti del formato nativo di sintassi. Le traduzioni sono descritte nei documenti [7], [8] presenti in fondo alla form di interrogazione.

L'OPAC viene interrogato direttamente nel proprio formato se non supporta più di una sintassi del record; in caso contrario viene presentato un menù a tendina che ne consente la scelta del formato.

Ogni OPAC fornisce gli indici di ricerca presenti nella pagina web di interrogazione in base agli *attributes use* del Bib-1 [9] che sup-

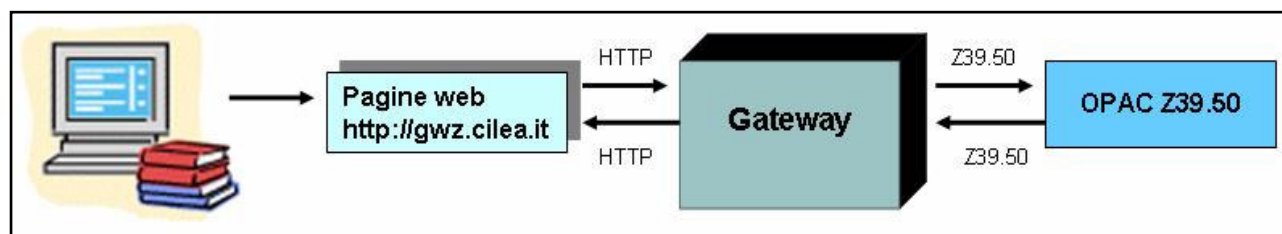


Fig 2 - Flusso delle operazioni del Gateway

guaggio PERL per la gestione interna.

La figura n. 2 mostra il flusso delle operazioni eseguite dal Gateway HTTP-Z39.50 durante l'interrogazione di un OPAC Z39.50.

porta. Queste pagine sono prodotte dal gestore con un programma apposito, in linguaggio PERL, che utilizza i dati OPAC raccolti in una banca dati interna sequenziale.

L'aggiornamento di questa banca dati viene fatta a cura del gestore mediante un caricamento automatico dopo aver verificato i dati.

Nella tabella dinamica all'URL:

<http://gwz.cilea.it/cgi-bin/reportOpac.cgi>

vengono visualizzati gli OPAC connessi al Gateway con i loro rispettivi attributi di caricamento.

Attualmente (dicembre 2005) sono stati caricati ventiquattro OPAC Italiani, che in maggioranza risultano utilizzare la sintassi UNIMARC. Sono noti al servizio gli attributi use, mentre le altre caratteristiche del retrieval dello standard non sono trattate esplicitamente, come pure si è scelto di usare come combinazione booleana tra le chiavi l'operatore AND. Ad una analisi iniziale, non approfondita, gli OPAC sembrano utilizzare profili differenti.

Il servizio Gateway e MAI, MetaOPAC Azalai Italiano

Dal 1999 è disponibile per la comunità di internet, ed in particolare la comunità bibliotecaria, il sistema MAI – MetaOPAC Azalai Italiano, descritto nel Bollettino del CILEA n. 68, Giugno 1999, e poi più ampiamente nel n. 79, Ottobre 2001.

Mediante la banca dati degli OPAC Italiani, MAI offre uno strumento per disporre nel Repertorio degli OPAC Italiani delle informazioni loro relative, mentre il software Azalai ne consente la metaricerca, cioè l'interrogazione cumulativa sincronizzata degli OPAC. Il modello di metaricerca realizzato dal sistema MAI è improntato alla interconnessione tra sistemi web, di cui si propone di mantenere e favorire la fruibilità tipica ipertestuale. L'OPAC web, ad esempio, fornisce attraverso la navigazione la visualizzazione del documento ed i dati di localizzazione alla biblioteca. Ulteriore eventuale vantaggio per l'utente è accedere, sempre attraverso la navigazione, ad altre funzioni che l'OPAC può fornire direttamente nel web, attuali e che possono essere prodotte solo localmente: ad esempio, la disponibilità del documento al prestito, prenotazione e consultazione.

Nella realtà dei cataloghi bibliografici italiani il protocollo Z39.50 risulta relativamente infrequente. Esistono però realizzazioni che danno disponibilità sia dell'interrogazione web, che l'uscita Z39.50. In alcuni altri casi l'OPAC è invece solamente di tipo Z39.50.

Il servizio di Gateway supporta la metaricerca MAI ponendosi come possibile interme-

diario di protocollo fisico verso web per la connessione delle realtà OPAC Z39.50, secondo quelle modalità sopra illustrate. Questi OPAC Z39.50, però, come abbiamo visto, attualmente non beneficiano ancora della navigabilità tra i dati e verso le funzioni. MAI perciò le interconnette preferibilmente nei casi di servizio OPAC in cui non è disponibile la versione web.

L'uscita web fornita dal Gateway, analogamente a quanto viene verificato per le applicazioni OPAC web native, va sottoposta all'analisi di compatibilità da parte del gestore MAI. In questi casi di OPAC sul Gateway, però, tecnicamente la compatibilità riguarda a grandi linee soltanto le condizioni poste al punto terzo della connettibilità MAI (ovvero il confronto positivo tra il profilo dell'installazione Z39.50 e l'information retrieval richiesto dal MAI). Il server Z39.50 farà poi fronte alle richieste MAI (applicabilità della connessione).

Attualmente sono stabilmente connesse via Gateway al MAI realizzazioni OPAC del software CyberPac di Geac (v. "Software per OPAC", [10]). Inoltre la connessione MAI a server Z39.50 realizzati con Cheshire per servizi OPAC Cilea è stata verificata positivamente sia dal punto di vista della compatibilità che dell'applicabilità (v. "Requisiti tecnici di connettibilità al MAI" [11]).

Il Gateway contribuisce così ad ampliare le possibilità e le modalità di connessione di cui questo progetto può disporre.

Bibliografia

- [1] Z39.50 Maintenance Agency Page, URL: <http://www.loc.gov/z3950/agency/>
- [2] HTTP, URL: <http://www.w3.org/Protocols/>
- [3] HTML, URL: <http://www.w3.org/TR/html401/>
- [4] ZAP!, URL: <http://www.indexdata.dk/zap/>
- [5] MAI - Metaopac Azalai Italiano, URL: <http://www.aib.it/aib/lis/opac1.htm>
- [6] URL: <http://httpd.apache.org/>
- [7] Etichettatura campi UNIMARC URL: <http://gwz.cilea.it/documentazione/EtichettaturaUNIMarc.html>
- [8] Etichettatura campi USMARC URL: <http://gwz.cilea.it/documentazione/EtichettaturaUSMARC.html>
- [9] Bib-1 Attribute set, URL: <http://www.loc.gov/z3950/agency/defs/bib1.html>
- [10] AIB: Associazione Italiana Biblioteche, Software per OPAC, URL: <http://www.aib.it/aib/opac/sw.htm>
- [11] URL: <http://www.aib.it/aib/opac/connect.htm>